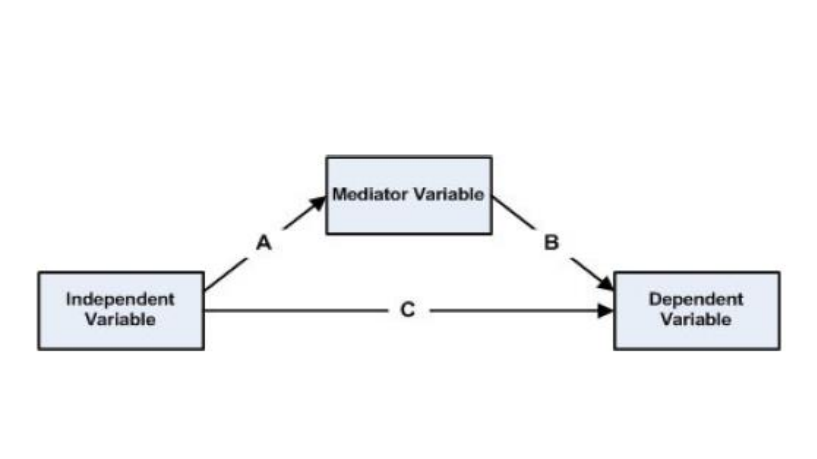
Mediation model for zero-inflated mediators

Zhigang Li
Department of Biostatistics
University of Florida

February 6, 2019

Introduction



Review

- Traditional mediation analysis, aka path analysis. Structural equation modeling. Judd and Kenny (1981), Baron and Kenny (1986), MacKinnon (2008)
- Potential-outcomes framework (ie, counterfactual-outcomes framework), aka. causal mediation analysis. VanderWeele (2009), Imai, Keele and Tingley (2010)
- Four assumptions for making causal conclusion

Motivation

- Zero-inflated data structure in microbiome data
- Disentangle the mediation effects due to presence of certain microbial taxa as well as the increase of abundance

Model for for Y

Let Y, M and X denote the outcome, mediator and certain exposure respectively. We construct a mediation model where Y has a distribution in exponential family with PDF or PMF given by:

$$f(y; \zeta, \psi) = \exp\Big((y\zeta - a(\zeta) + b(y))/\psi\Big).$$

By this formulation, we have $E(Y) = a'(\zeta)$ and var$(Y) = a''(\zeta)\psi$. We first construct the association of Y with M and X using GLM (McCullagh and Nelder,1989):

$$g(E(Y)) = \beta_0 + \beta_1 M + \beta_2 1_{(M>0)} + \beta_3 X$$

where $g(\cdot)$ is a known link function and $1_{(\cdot)}$ is an indicator function. Under this model, the parameters ζ and ψ can considered as functions of β. For example, for normal distribution and identity link function, we have the mean parameter $\zeta = \beta_0 + \beta_1 M + \beta_2 1_{(M>0)} + \beta_3 X$ and the variance parameter ψ is a constant function of β where $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)$. Hereafter, we write $\zeta(\beta)$ and $\psi(\beta)$.

Zero-inflated mediators

Let M denote the mediator with mixture distribution with the two-part density function given by:

$$f(m;\theta) = \begin{cases} \mathcal{G}(\theta), & m = 0 \\ (1 - \mathcal{G}(\theta))G(m;\theta), & m > 0 \end{cases}$$

where θ is the K-dimensional parameter vector associated with the distribution, $0 < \mathcal{G}(\theta) < 1$ is the probability of M taking value of 0 and $G(m;\theta)$ is the conditional density (or mass) function of M given that M is positive. Examples include zero-inflated log-normal (ZILoN) distribution, zero-inflated beta distribution (ZIB) and zero-inflated Poisson distribution (ZIP).

Examples: zero-inflated log-normal distribution

If the meditor M has a ZILoN distribution, we have

$$f(m;\theta) = \begin{cases} \Delta, & m = 0 \\ (1-\Delta)G(m;\theta), & m > 0 \end{cases}$$

where $\theta = (\mu, \sigma, \Delta)$, $\mathcal{G}(\theta) = \Delta$ and

$$G(m;\theta) = \frac{1}{m\sigma\sqrt{2\pi}} \exp\left(- \frac{(\log(m) - \mu)^2}{2\sigma^2} \right).$$

Examples: zero-inflated beta distribution

If the mediator M has a ZIB distribution, we have

$$f(m;\theta) = \begin{cases} \Delta, & m = 0 \\ (1-\Delta)G(m;\theta), & m > 0 \end{cases}$$

where $\theta = (\mu, \phi, \Delta)$, $\mathcal{G}(\theta) = \Delta$ and

$$G(m;\theta) = \frac{m^{\mu\phi-1}(1-m)^{(1-\mu)\phi-1}}{B(\mu\phi,(1-\mu)\phi)}, \quad m \in (0,1), 1 > \mu > 0, \phi > 0.$$

Examples: zero-inflated Poisson distribution

For ZIP mediator, the two-part density function is

$$f(m) = \begin{cases} \mathcal{G}(\theta), & m = 0 \\ (1 - \mathcal{G}(\theta))G(m; \theta), & m = 1, 2, \ldots \end{cases}$$

where $\theta = (\lambda, \Delta^*)$ and

$$\mathcal{G}(\theta) = \Delta^* + (1 - \Delta^*) \exp(-\lambda)$$

$$G(m; \theta) = \frac{\lambda^m}{m!(\exp(\lambda) - 1)}, \quad m = 1, 2, \ldots.$$

Here λ is the Poisson parameter and $0 < \Delta^* < 1$ is the parameter controlling the excessive zeros.

General mediation model

We assume the following equations for the mediation model.

$$g(E(Y)) = \beta_0 + \beta_1 M + \beta_2 1_{(M>0)} + \beta_3 X$$
$$T(\theta) = \nu_0 + \nu_1 X$$

where $T : R^K \to R^K$ is a known one-to-one (possibly nonlinear) transformatin of the paramter vector θ, and ν_0 and ν_1 are two K-dimensional regressional coefficients where ν_0 can be considered as the intercept vector and ν_1 as the slope vector.

Mediation model: continuous outcome and ZILoN mediator

A mediation model for a continuous Y and ZILoN M consists of the following equations:

$$
\begin{aligned}
&Y = \beta_0 + \beta_1 M + \beta_2 1_{(M>0)} + \beta_3 X + \epsilon \\
&\mu = \alpha_0 + \alpha_1 X \\
&\log(\sigma) = \xi_0 + \xi_1 X \\
&\log\left(\frac{\Delta}{1-\Delta}\right) = \gamma_0 + \gamma_1 X
\end{aligned}
$$

where ϵ follows a normal distribution with mean of zero and variance δ^2.

Mediation model: binary outcome and ZILoN mediator

A mediation model for a binary Y and ZILoN M consists of the following equations:

$$\log\left(\frac{P(Y=1)}{1-P(Y=1)}\right) = \beta_0 + \beta_1 M + \beta_2 1_{(M>0)} + \beta_3 X$$

$$\mu = \alpha_0 + \alpha_1 X$$

$$\log(\sigma) = \xi_0 + \xi_1 X$$

$$\log\left(\frac{\Delta}{1-\Delta}\right) = \gamma_0 + \gamma_1 X$$

Natural Indirect Effect (NIE), Natural Direct Effect (NDE), Controlled Direct Effect (CDE)

The NIE, NDE and CDE for X changing from x_1 to x_2 are defined as:

$$\text{NIE} = E(Y_{x_2 M_{x_2}} - Y_{x_2 M_{x_1}})$$
$$\text{NDE} = E(Y_{x_2 M_{x_1}} - Y_{x_1 M_{x_1}})$$
$$\text{CDE} = E(Y_{x_2 m} - Y_{x_1 m}), \text{ for a fixed (i.e., controlled) m.}$$

Natural Indirect Effect

By using Riemann-Stieltjes integral:

$$\begin{aligned}\text{NIE} &= E(Y_{x_2 M_{x_2}}) - E(Y_{x_2 M_{x_1}}) = E(E(Y_{x_2 M_{x_2}}|M_{x_2})) - E(E(Y_{x_2 M_{x_1}}|M_{x_1}))\\ &= \int\limits_{m\in\Omega} E(Y_{x_2 M_{x_2}}|M_{x_2}=m)dF_{M_{x_2}}(m) - \int\limits_{m\in\Omega} E(Y_{x_2 M_{x_1}}|M_{x_1}=m)dF_{M_{x_1}}(m)\\ &= \text{NIE}_1 + \text{NIE}_2,\end{aligned}$$

where

$$\begin{aligned}\text{NIE}_1 &= \int\limits_{m\in\Omega\backslash 0} g^{-1}(\beta_0+\beta_1 m+\beta_2+\beta_3 x_2)G(m;\theta_{x_2})dm\\ &\quad - \int\limits_{m\in\Omega\backslash 0} g^{-1}(\beta_0+\beta_1 m+\beta_2+\beta_3 x_2)G(m;\theta_{x_1})dm,\\ \text{NIE}_2 &= \Delta_{x_2}\left(g^{-1}(\beta_0+\beta_3 x_2) - \int\limits_{m\in\Omega\backslash 0} g^{-1}(\beta_0+\beta_1 m+\beta_2+\beta_3 x_2)G(m;\theta_{x_2})dm\right)\\ &\quad - \Delta_{x_1}\left(g^{-1}(\beta_0+\beta_3 x_2) - \int\limits_{m\in\Omega\backslash 0} g^{-1}(\beta_0+\beta_1 m+\beta_2+\beta_3 x_2)G(m;\theta_{x_1})dm\right)\end{aligned}$$

Natural Direct Effect

$$\begin{aligned}\text{NDE} &= E\Big(Y_{x_2 M_{x_1}} - Y_{x_1 M_{x_1}}\Big) = E\Big(Y_{x_2 M_{x_1}}\Big) - E\Big(Y_{x_1 M_{x_1}}\Big)\\ &= \text{NDE}_1 + \text{NDE}_2,\end{aligned}$$

where

$$\begin{aligned}\text{NDE}_1 = &\int_{m \in \Omega \setminus 0} g^{-1}(\beta_0 + \beta_1 m + \beta_2 + \beta_3 x_2))G(m;\theta_{x_1})dm\\ &- \int_{m \in \Omega \setminus 0} g^{-1}(\beta_0 + \beta_1 m + \beta_2 + \beta_3 x_1)G(m;\theta_{x_1})dm,\end{aligned}$$

$$\begin{aligned}\text{NDE}_2 = \Delta_{x_1}\Bigg(&g^{-1}(\beta_0 + \beta_3 x_2) - \int_{m \in \Omega \setminus 0} g^{-1}(\beta_0 + \beta_1 m + \beta_2 + \beta_3 x_2))G(m;\theta_{x_1})dm\\ &- g^{-1}(\beta_0 + \beta_3 x_1) + \int_{m \in \Omega \setminus 0} g^{-1}(\beta_0 + \beta_1 m + \beta_2 + \beta_3 x_1)G(m;\theta_{x_1})dm\Bigg).\end{aligned}$$

Controlled Direct Effect

$$\text{CDE} = E\Big(Y_{x_2 m} - Y_{x_1 m}\Big) = E\Big(Y_{x_2 m}\Big) - E\Big(Y_{x_1 m}\Big)$$
$$= g^{-1}(\beta_0 + \beta_1 m + \beta_2 1_{(m>0)} + \beta_3 x_2) - g^{-1}(\beta_0 + \beta_1 m + \beta_2 1_{(m>0)} + \beta_3 x_1).$$

When $g(\cdot)$ is identity link, CDE $= \beta_3(x_2 - x_1)$.

Mechanisms for observing zero's

Let M^* denote the observed value of the mediator M. When the observed value is positive (i.e., $M^* > 0$), we know that $M^* = M$. But when $M^* = 0$, we don't know whether M is truly zero or M is posively but observed as zero. A possible mechanisms for observing zero could be:

$$P(M^* = 0|M) = 1_{(M<c)}$$

where c denotes the limit of detection (LOD). This model requires that all true positive values below LOD have an observed value of zero which is commonly seen in epidemiologic studies. The specification of the model for observing zero can be flexible. Another example could be:

$$P(M^* = 0|M) = \exp(-\eta M)$$

where $\eta > 0$ and thus it is a decreasing function of M such that smaller values of M are more likely to be observed as zero than larger values. Notice that the observed value M^* is equal to zero with probability of one when $M = 0$ which corresponds to the case that M is truly zero. **We assume that the probability of observing a zero only depends on M which also means that it is independent of Y conditional on M.**

Maximum Likelihood Estimator (MLE)

The observed data for each subject can be denoted by this vector (Y, M^*, R, X) where $R = 1_{(M^*>0)}$. The first group consists of subjects whose observed values of M are non-zero (i.e., $M^* > 0$ and $M^* = M$) and the log-likelihood contribution for the ith subject can be calculated as:

$$
\begin{aligned}
l_i^1 &= \log(f(y_i, r_i|m_i, x_i)f(m_i|x_i)) = \log(f(y_i|m_i, x_i)f(r_i|m_i, x_i)f(m_i|x_i)) \\
&= \log(f(y_i|m_i, x_i)) + \log(f(r_i|m_i)) + \log(f(m_i|x_i)) \\
&= \log(f(y_i; \zeta(\beta), \psi(\beta))) + \log(1 - P(M^* = 0|M = m_i)) + \log((1 - \Delta_{x_i})G(m_i; \theta_{x_i})).
\end{aligned}
$$

Let $F(m|x)$ denote the (conditional) cumulative distribution function for M. The second group consists of subjects whose observed M are 0 (ie, $M^* = 0$) and the log-likelihood contribution can be calculated as:

$$
l_i^2 = \log(f(y_i, r_i, m_i|x_i)) = \log\left(\int_{m\in\Omega} f(y_i|m, x_i)f(r_i|m)dF(m|x_i)\right)
$$

$$
= \log\Bigg(\Delta_{x_i} f(y_i; \zeta(\beta), \psi(\beta))
$$

$$
+ (1 - \Delta_{x_i}) \int_{m\in\Omega\backslash 0} f(y_i|m, x_i)P(M^* = 0|M = m))G(m; \theta_{x_i})dm.\Bigg)
$$

Taken together, we have the complete log-likelihood function is given by

$$
l = \sum_{i\in\text{group 1}} l_i^1 + \sum_{i\in\text{group 2}} l_i^2.
$$

Closed-form expression for l can be derived for specific case although there is no general closed-form expression.

Simulation

100 data sets, $n = 300$ and X is generated by $N(0,1)$.

Parameter /Effect	True	Probability case: Mean Estimate	Bias	Bias %	SE	Mean SE	CP(%)	LOD case: Mean Estimate	Bias	Bias %	SE	Mean SE	CP(%)
						Our approach							
NIE1	0.088	0.123	-0.035	-39.8	0.01	0.14	94	0.138	-0.050	-56.8	0.02	0.15	95
NIE2	-0.154	-0.114	-0.040	26.0	0.02	0.20	96	-0.118	-0.036	23.4	0.02	0.21	98
NIE	-0.066	0.015	-0.081	122.7	0.03	0.27	96	0.020	-0.086	130.3	0.03	0.29	96
NDE	1.000	0.960	0.040	4.0	0.01	0.07	94	0.970	0.030	3.0	0.01	0.07	94
CDE	1.000	0.960	0.040	4.0	0.01	0.07	94	0.970	0.030	3.0	0.01	0.07	94
β_0	1.000	2.267	-1.267	-126.7	0.07	0.27	100	1.944	-0.944	-94.4	0.07	0.23	100
β_1	1.000	1.103	-0.103	-10.3	0.01	0.04	100	1.074	-0.074	-7.4	0.01	0.04	100
β_2	10.000	8.295	1.705	17.1	0.08	0.34	100	8.756	1.244	12.4	0.08	0.28	100
β_3	1.000	0.960	0.04	4.0	0.01	0.07	100	0.970	0.030	3.0	0.01	0.07	100
α_0	-0.500	-0.417	-0.083	16.6	0.01	0.10	100	-0.454	-0.046	9.2	0.01	0.08	100
α_1	0.100	0.110	-0.010	-10.0	0.01	0.09	100	0.105	-0.005	-5.0	0.01	0.08	100
ξ_0	0.200	0.134	0.066	33.0	0.01	0.07	100	0.181	0.019	9.5	0.01	0.06	100
ξ_1	0.000	0.003	-0.003	-	0.01	0.07	100	0.009	-0.009	-	0.01	0.06	100
γ_0	-1.500	-1.492	-0.008	0.5	0.01	0.15	100	-1.492	-0.008	0.5	0.01	0.15	100
γ_1	0.100	0.082	0.018	18.0	0.01	0.15	100	0.082	0.018	18.0	0.01	0.15	100
δ	1.000	1.195	-0.195	-19.5	0.02	0.09	100	1.126	-0.126	-12.6	0.01	0.07	99
η	1.300	1.254	0.046	3.5	0.01	0.08	100						
						Imai, Keele and Tingley (2010)							
ACME(NIE)	-0.066	0.120	-0.186	281.8	-	-	69	0.119	-0.185	280.3	-	-	69
ADE(NDE)	1.000	0.813	0.187	18.7	-	-	90	0.799	0.201	20.1	-	-	85

Table: Results for ZILoN M and continuous Y.

Simulation

Parameter /Effect	True	Probability case: Mean Estimate	Bias	Bias %	SE	Mean SE	CP(%)	LOD case: Mean Estimate	Bias	Bias %	SE	Mean SE	CP(%)
						Our approach							
OR^{NIE}	0.948	0.953	-0.005	-0.5	0.01	0.14	85	0.970	-0.022	-2.3	0.01	0.12	88
OR^{NDE}	2.334	2.300	0.034	1.5	0.04	0.37	91	2.284	0.050	2.1	0.03	0.36	94
OR^{CDE}	2.718	2.823	-0.105	-3.9	0.07	0.62	91	2.749	-0.031	-1.1	0.06	0.55	95
RDNIE	-0.010	-0.010	0.000	0.0	-	-	87	-0.007	-0.003	30.0	-	-	89
RDNDE	0.181	0.173	0.008	4.4	-	-	92	0.172	0.009	5.0	-	-	93
β_0	-1.000	-1.121	0.121	-12.1	0.08	0.85	100	-1.059	0.059	-5.9	0.08	0.70	100
β_1	1.000	1.068	-0.068	-6.8	0.03	0.32	100	1.084	-0.084	-8.4	0.03	0.31	100
β_2	1.000	1.133	-0.133	-13.3	0.10	0.98	100	1.044	-0.044	-4.4	0.09	0.85	100
β_3	1.000	1.013	-0.013	-1.3	0.02	0.21	100	0.990	0.010	1.0	0.02	0.20	100
α_0	-1.000	-0.896	-0.104	10.4	0.02	0.26	100	-0.912	-0.088	8.8	0.02	0.22	100
α_1	0.500	0.510	-0.010	-2.0	0.02	0.18	100	0.463	0.037	7.4	0.02	0.17	100
ξ_0	0.500	0.460	0.040	8.0	0.01	0.08	100	0.456	0.044	8.8	0.01	0.09	100
ξ_1	0.000	-0.011	0.011	-	0.01	0.08	100	0.009	-0.009	-	0.01	0.08	100
γ_0	-1.500	-1.378	-0.122	8.1	0.05	0.81	100	-1.429	-0.071	4.7	0.04	0.52	100
γ_1	0.500	0.477	0.023	4.6	0.04	0.41	100	0.426	0.074	14.8	0.04	0.36	100
η	2.000	2.122	-0.122	-6.1	0.03	0.46	100						
						Imai, Keele and Tingley (2010)							
ACME(RDNIE)	-0.010	0.032	-0.042	420.0	-	-	31	0.032	-0.042	420.0	-	-	30
ADE(RDNDE)	0.181	0.086	0.095	52.5	-	-	18	0.086	0.095	52.5	-	-	13

Table: Results for ZILoN M and binary Y.

Simulation

Parameter /Effect	True	Probability case: Mean Estimate	Bias	Bias %	SE	Mean SE	CP(%)	LOD case: Mean Estimate	Bias	Bias %	SE	Mean SE	CP(%)
						Our approach							
NIE1	-0.013	-0.020	0.007	-53.8	0.00	0.02	93	-0.017	0.004	-30.8	0.00	0.02	96
NIE2	-0.632	-0.726	0.094	-14.9	0.03	0.26	96	-0.735	0.103	-16.3	0.03	0.26	94
NIE	-0.645	-0.746	0.101	-15.7	0.03	0.28	94	-0.752	0.107	-16.6	0.03	0.28	94
NDE	1.000	0.989	0.011	1.1	0.01	0.06	91	0.995	0.005	0.5	0.01	0.06	94
CDE	1.000	0.989	0.011	1.1	0.01	0.06	91	0.995	0.005	0.5	0.01	0.06	94
β_0	1.000	1.058	-0.058	-5.8	0.02	0.17	100	0.995	0.005	0.5	0.02	0.16	100
β_1	1.000	1.137	-0.137	-13.7	0.06	0.43	100	0.982	0.018	1.8	0.04	0.36	100
β_2	10.000	9.854	0.146	1.5	0.05	0.29	100	10.009	-0.009	-0.1	0.03	0.26	100
β_3	1.000	0.989	0.011	1.1	0.01	0.06	100	0.995	0.005	0.5	0.01	0.06	100
α_0	0.100	0.110	-0.010	-10.0	0.01	0.05	100	0.103	-0.003	-3.0	0.00	0.04	100
α_1	0.100	0.102	-0.002	-2.0	0.01	0.05	100	0.106	-0.006	-6.0	0.00	0.04	100
ξ_0	2.000	2.043	-0.043	-2.2	0.01	0.11	100	2.020	-0.020	-1.0	0.01	0.08	100
ξ_1	0.100	0.095	0.005	5.0	0.01	0.12	100	0.106	-0.006	-6.0	0.01	0.09	100
γ_0	-2.000	-2.054	0.054	-2.7	0.02	0.19	100	-2.054	0.054	-2.7	0.02	0.19	100
γ_1	0.500	0.580	-0.080	-16.0	0.02	0.19	100	0.580	-0.080	-16.0	0.02	0.19	100
δ	1.000	0.994	0.006	0.6	0.01	0.04	100	0.994	0.006	0.6	0.00	0.04	100
η	1.500	1.495	0.005	0.3	0.00	0.07	100						
						Imai, Keele and Tingley (2010)							
ACME(NIE)	-0.645	-0.001	-0.644	99.8	-	-	0	-0.009	-0.636	98.6	-	-	0
ADE(NDE)	1.000	0.429	0.571	57.1	-	-	15	0.790	0.210	21.0	-	-	61

Table: Results for ZIB M and continuous Y.

Simulation

Parameter /Effect	True	Probability case: Mean Estimate	Bias	Bias %	SE	Mean SE	CP(%)	LOD case: Mean Estimate	Bias	Bias %	SE	Mean SE	CP(%)
						Our approach							
OR^{NIE}	0.820	0.809	0.011	1.3	0.01	0.10	94	0.809	0.011	1.3	0.01	0.09	98
OR^{NDE}	2.528	2.653	-0.125	-4.9	0.05	0.46	94	2.646	-0.118	-4.7	0.05	0.45	96
OR^{CDE}	2.718	2.926	-0.208	-7.7	0.06	0.59	94	2.967	-0.249	-9.2	0.06	0.59	97
RDNIE	-0.040	-0.042	0.002	-5.0	-	-	94	-0.042	0.002	-5.0	-	-	98
RDNDE	0.209	0.212	-0.003	-1.4	-	-	94	0.212	-0.003	-1.4	-	-	97
β_0	-1.000	-1.091	0.091	-9.1	0.12	0.97	100	-1.283	0.283	-28.3	0.13	1.38	100
β_1	1.000	0.996	0.004	0.4	0.11	1.07	100	1.073	-0.073	-7.3	0.14	1.23	100
β_2	1.000	1.114	-0.114	-11.4	0.14	1.18	100	1.280	-0.280	-28.0	0.16	1.76	100
β_3	1.000	1.054	-0.054	-5.4	0.02	0.20	100	1.069	-0.069	-6.9	0.02	0.19	100
α_0	-0.500	-0.451	-0.049	9.8	0.01	0.08	100	-0.474	-0.026	5.2	0.01	0.08	100
α_1	-0.500	-0.474	-0.026	5.2	0.01	0.05	100	-0.486	-0.014	2.8	0.01	0.05	100
ξ_0	2.000	2.043	-0.043	-2.2	0.01	0.12	100	2.043	-0.043	-2.2	0.01	0.13	100
ξ_1	-1.000	-0.984	-0.016	1.6	0.01	0.11	100	-0.980	-0.020	2.0	0.01	0.12	100
γ_0	-1.500	-1.366	-0.134	8.9	0.08	0.75	100	-1.492	-0.008	0.5	0.04	0.37	100
γ_1	0.500	0.657	-0.157	-31.4	0.05	0.38	100	0.534	-0.034	-6.8	0.03	0.32	100
η	2'000	2.353	-0.353	-17.7	0.06	0.64	100						
						Imai, Keele and Tingley (2010)							
ACME(RDNIE)	-0.040	-0.017	-0.023	57.5	-	-	80	-0.025	-0.015	37.5	-	-	86
ADE(RDNDE)	0.209	0.205	0.004	1.9	-	-	98	0.219	-0.01	-4.8	-	-	97

Table: Results for ZIB M and binary Y.

Simulation

Parameter /Effect	True	Probability case Mean Estimate	Bias	Bias %	SE	Mean SE	CP(%)	LOD case Mean Estimate	Bias	Bias %	SE	Mean SE	CP(%)
						Our approach							
NIE1	0.004	0.054	-0.050	-1250.0	0.01	0.04	74	0.013	-0.009	-225.0	0.00	0.01	93
NIE2	0.055	0.043	0.012	21.8	0.00	0.02	82	0.065	-0.010	-18.2	0.00	0.02	86
NIE	0.059	0.097	-0.038	-64.4	0.01	0.05	88	0.078	-0.019	-32.2	0.00	0.03	87
NDE	1.000	0.941	0.059	5.9	0.01	0.07	88	0.988	0.012	1.2	0.01	0.06	95
CDE	1.000	0.941	0.059	5.9	0.01	0.07	88	0.988	0.012	1.2	0.01	0.06	95
β_0	1.000	1.800	-0.800	-80.0	0.07	0.23	99	1.203	-0.203	-20.3	0.04	0.17	100
β_1	0.010	0.167	-0.157	-1570.0	0.01	0.04	99	0.036	-0.026	-260.0	0.00	0.03	100
β_2	15.000	13.400	1.600	10.7	0.12	0.37	99	14.650	0.350	2.3	0.06	0.22	100
β_3	1.000	0.941	0.059	5.9	0.01	0.07	100	0.988	0.012	1.2	0.01	0.06	100
α_0	1.500	1.517	-0.017	-1.1	0.00	0.04	100	1.462	0.038	2.5	0.00	0.03	100
α_1	0.100	0.091	0.009	9.0	0.00	0.04	100	0.105	-0.005	-5.0	0.00	0.03	100
γ_0	-2.000	-2.043	0.043	-2.2	0.02	0.2	100	-2.077	0.077	-3.9	0.02	0.2	100
γ_1	0.100	0.088	0.012	12.0	0.02	0.20	100	0.104	-0.004	-4.0	0.02	0.20	100
δ	1.000	1.087	-0.087	-8.7	0.00	0.02	99	1.011	-0.011	-1.1	0.01	0.04	100
η	0.500	0.482	0.018	3.6	0.01	0.06	100						
						Imai, Keele and Tingley (2010)							
ACME(NIE)	0.059	0.010	0.049	83.1	-	-	49	0.020	0.039	66.1	-	-	61
ADE(NDE)	1.000	0.826	0.174	17.4	-	-	89	0.668	0.332	33.2	-	-	71

Table: Results for ZIP M and continuous Y.

Simulation

Parameter /Effect	True	Probability case Mean Estimate	Bias	Bias %	SE	Mean SE	CP(%)	LOD case Mean Estimate	Bias	Bias %	SE	Mean SE	CP(%)
						Our approach							
OR^{NIE}	5.005	7.177	-2.172	-43.4	0.58	3.56	99	4.217	0.788	15.7	0.08	0.74	64
OR^{NDE}	1.818	1.787	0.031	1.7	0.04	0.4	94	2.072	-0.254	-14.0	0.04	0.41	98
OR^{CDE}	2.718	2.714	0.004	0.1	0.08	0.85	94	3.039	-0.321	-11.8	0.08	0.83	99
RDNIE	0.359	0.388	-0.029	-8.1	-	-	93	0.322	0.037	10.3	-	-	82
RDNDE	0.081	0.069	0.012	14.8	-	-	89	0.103	-0.022	-27.2	-	-	92
β_0	-0.500	-0.440	-0.060	12.0	0.03	0.35	100	-0.697	0.197	-39.4	0.03	0.27	98
β_1	1.000	1.026	-0.026	-2.6	0.02	0.16	100	1.050	-0.050	-5.0	0.02	0.16	100
β_2	-9 .000	-9.299	0.299	-3.3	0.14	1.53	100	-9.210	0.210	-2.3	0.16	1.51	100
β_3	1.000	0.957	0.043	4.3	0.03	0.31	100	1.083	-0.083	-8.3	0.02	0.27	100
α_0	2.000	1.990	0.010	0.5	0.00	0.04	100	2.035	-0.035	-1.8	0.00	0.04	100
α_1	-0.500	-0.505	0.005	-1.0	0.00	0.03	100	-0.477	-0.023	4.6	0.00	0.03	100
γ_0	-0.700	-0.916	0.216	-30.9	0.04	0.43	100	-0.436	-0.264	37.7	0.02	0.2	99
γ_1	3.000	3.394	-0.394	-13.1	0.08	0.64	100	3.194	-0.194	-6.5	0.05	0.42	100
η	0.600	0.571	0.029	4.8	0.01	0.09	100						
						Imai, Keele and Tingley (2010)							
ACME(RDNIE)	0.359	-0.292	0.651	181.3	-	-	0	-0.098	0.457	127.3	-	-	0
ADE(RDNDE)	0.081	0.053	0.028	34.6	-	-	53	0.087	-0.006	-7.4	-	-	96

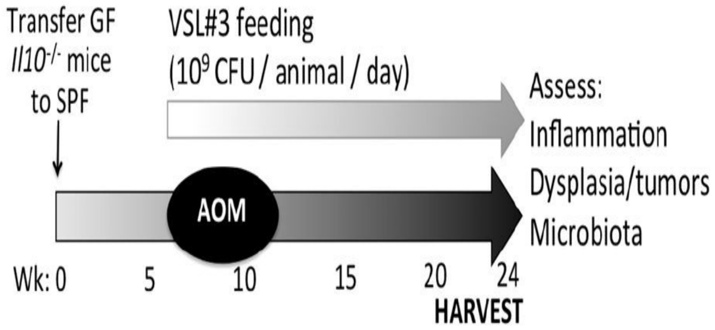
Table: Results for ZIP M and binary Y.

Real study

- To assess the ability of probiotic cocktail VSL#3 to alter the colonic microbiota and decrease inflammation-associated colorectal cancer when administered as interventional therapy after the onset of inflammation. Arthur et al. (2013).
- VSL#3 is a commercially available probiotic cocktail (Sigma-Tau Pharmaceuticals, Inc.) of eight strains of lactic acid-producing bacteria: Lactobacillus plantarum, Lactobacillus delbrueckii subsp. Bulgaricus, Lactobacillus paracasei, Lactobacillus acidophilus, Bifidobacterium breve, Bifidobacterium longum, Bifidobacterium infantis, and Streptococcus salivarius subsp.
- Orally administered VSL#3 has shown success in ameliorating symptoms and reducing inflammation in human pouchitis and ulcerative colitis. Preventive VSL#3 administration can also attenuate colitis in Il10-/- mice and ileitis in SAMP1/YitFc mice. VSL#3 probiotic has the capability to prevent inflammation and carcinogenesis when used as a preventative strategy.
- However, the therapeutic potential of VSL#3 probiotic administered after the onset of inflammation remains unknown.
- 33 mice, 19 treated VS 14 control. Microbiome data: 16S sequencing on stool samples. Outcome variable: dysplasia score.

Real study

- Germ free (GF) Il10-/- mice were transferred to specific pathogen free (SPF) conditions where they were colonized with SPF commensal bacteria for 7 weeks to establish intestinal inflammation.Conventionalizing mice born and raised in GF conditions negates the confounding factor of familial transmission of the microbiota.
- Tumorigenesis was then initiated with 6 weekly injections of Azoxymethane (AOM, 10 mg kg), and VSL#3 was orally administered (109 CFU/mouse/day) to mice daily throughout the remainder of the experiment. After 24 weeks, mice were sacrificed and tissues were harvested for assessment of inflammation, and tumorigenesis.



Real study

OTU	Zero-inflated log-normal				Zero-inflated beta			
	NIE1	CI	NIE2	CI	NIE1	CI	NIE2	CI
1	0.004	(-0.127, 0.134)	0.015	(-0.476,0.506)	-0.007	(-0.141,0.127)	0.000	(-0.009,0.009)
2	0.034	(-0.297, 0.365)	-0.077	(-0.660, 0.506)	0.005	(-0.069,0.080)	-000	(-0.012,0.012)
3	0.202	(-0.686,1.091)	-0.343	(-0.941, 0.255)	0.250	(-0.293,0.794)	0.000	(-0.008,0.008)
4	-0.010	(-0.133, 0.111)	0.474	(-0.264, 1.211)	0.013	(-0.076,0.102)	0.470	(-0.268,1.209)
5	-0.232	(-1.607, 1.144)	0.015	(-0.263, 0.292)	-0.060	(-0.309,0.189)	0.000	(-0.007,0.007)
6	0.196	(-2.917,3.309)	-0.303	(-0.921, 0.315)	0.004	(-0.156,0.164)	0.000	$(-1.4\times10^{-10}, 1.3\times10^{-10})$
7	0.229	(-4.512, 4.969)	-0.233	(-0.810, 0.343)	0.024	(-0.219,0.266)	-0.234	(-0.806,0.339)
8	0.004	(-0.120, 0.127)	-0.372	(-1.009,0.265)	0.006	(-0.194,0.206)	-0.366	(-1.005,0.273)
9	0.001	(-0.080,0.082)	1.656	(0.619,2.694)	0.000	(-0.073,0.074)	0.000	(-0.034,0.034)
10	-0.004	(-0.177, 0.169)	0.051	(-0.493,0.596)	-0.006	(-0.100,0.089)	0.000	$(-2.5\times10^{-04}, 2.4\times10^{-4})$
11	0.246	(-0.457,0.948)	-0.009	(-0.354,0.337)	0.094	(-0.295,0.484)	-0.411	(-1.083,0.260)
12	3.422	(-107.995,114.840)	0.000	$(-1.7\times10^{-03}, 1.7\times10^{-03})$	-0.061	(-0.297,0.175)	0.002	(-0.069,0.074)
13	-0.019	(-0.163, 0.124)	-0.302	(-0.918, 0.314)	-0.040	(-0.269,0.190)	-0.292	(-0.909,0.324)
14	0.213	(-4.425, 4.852)	0.462	(-0.265, 1.190)	0.001	(-0.152,0.154)	0.000	(-0.009,0.009)
15	0.098	(-0.409,0.604)	-0.286	(-0.868, 0.295)	0.139	(-0.236,0.513)	0.000	$(-3.9\times10^{-12}, 3.6\times10^{-12})$
16	0.409	(-2.254,3.072)	-0.163	(-1.340, 1.014)	0.089	(-0.224,0.402)	0.000	(-0.004,0.004)
17	-0.006	(-0.131, 0.118)	0.070	(-0.285, 0.426)	-0.009	(-0.318,0.300)	-0.001	(-0.016,0.015)
18	1.162	(-13.871,16.195)	-0.322	(-0.943, 0.300)	0.032	(-0.185,0.248)	-0.306	(-0.929,0.318)
19	-0.095	(-1.123, 0.933)	0.225	(-0.326, 0.776)	-0.052	(-0.302,0.197)	0.000	$(-2.5\times10^{-06}, 2.7\times10^{-06})$
20	0.009	(-0.220,0.239)	0.083	(-0.921, 1.087)	0.005	(-0.252,0.262)	0.341	(-0.326,1.009)
21	0.349	(-2.533,3.230)	-0.147	(-1.378,1.084)	0.059	(-0.199,0.318)	2.1×10^{-109}	$(2.1\times10^{-109}, 2.2\times10^{-109})$
22	0.007	(-0.421, 0.435)	-0.286	(-0.923,0.350)	0.000	(-0.029,0.028)	0.000	$(-1.4\times10^{-06}, 1.3\times10^{-06})$
23	0.232	(-0.804,1.268)	-0.005	(-1.347, 1.337)	0.127	(-0.285,0.540)	0.000	(-0.002,0.002)
24	0.055	(-0.352,0.462)	-0.028	(-0.377,0.321)	0.026	(-0.180,0.231)	-0.315	(-0.994,0.364)
25	0.103	(-0.863,1.069)	0.232	(-0.322,0.786)	0.048	(-0.216,0.312)	0.232	(-0.322,0.786)
26	-0.170	(-1.027,0.687)	0.141	(-0.310, 0.593)	-0.107	(-0.445,0.231)	0.007	(-0.262,0.276)
27	-0.007	(-0.688, 0.674)	0.567	(-0.209, 1.344)	0.003	(-0.279,0.286)	0.570	(-0.204, 1.345)
28	-0.079	(-1.309, 1.151)	-0.605	(-1.463, 0.253)	0.008	(-0.141,0.157)	-0.597	(-1.451,0.257)
29	-0.003	(-0.169,0.164)	0.383	(-0.308, 1.075)	-0.001	(-0.075,0.072)	0.383	(-0.308,1.075)
30	-0.003	(-0.417, 0.412)	0.076	(-0.340, 0.493)	0.001	(-0.106,0.108)	0.000	$(-1.6\times10^{-05}, 1.9\times10^{-05})$
31	0.228	(-1.022, 1.477)	0.020	(-0.262, 0.301)	0.125	(-0.256,0.506)	0.147	(-0.336,0.629)
32	-1.6×10^{05}	$(-6.2\times10^{06}, 5.9\times10^{06})$	0.294	(-0.323, 0.910)	-0.005	(-0.102,0.092)	0.000	$(-1.6\times10^{-14}, 1.5^{-14})$
33	0.226	(-0.875, 1.327)	0.070	(-0.239,0.380)	0.128	(-0.235,0.490)	0.000	$(-7.9\times10^{-11}, 7.5\times10^{-11})$
34	0.112	(-0.620, 0.844)	-0.002	(-0.079,0.075)	0.070	(-0.279,0.419)	-0.152	(-0.794,0.489)
35	0.036	(-0.297,0.368)	-0.040	(-0.857,0.776)	0.051	(-0.199,0.300)	0.383	(-0.308,1.073)
36	-0.090	(-0.645,0.465)	-0.362	(-0.990,0.265)	-0.004	(-0.105,0.098)	0.000	(-0.026,0.024)

Real study

Zero-inflated beta										
	Our approach						Imai, Keele and Tingley (2010)			
OTU	NIE	CI	NDE	CI	CDE	CI	ACME(NIE)	CI	ADE(NDE)	CI
1	-0.007	(-0.141,0.127)	1.660	(0.615, 2.705)	1.660	(0.615, 2.705)	0.000	(-0.166, 0.170)	1.690	(0.602, 2.770)
2	0.005	(-0.071,0.080)	1.649	(0.608,2.689)	1.649	(0.608,2.689)	0.003	(-0.197, 0.231)	1.623	(0.773, 2.335)
3	0.250	(-0.294, 0.794)	1.391	(0.352,2.429)	1.391	(0.352,2.429)	0.218	(-0.231, 1.000)	1.418	(0.427, 2.502)
4	0.484	(-0.255,1.221)	1.176	(0.419,1.933)	1.176	(0.419,1.933)	0.008	(-0.290, 0.489)	1.667	(0.737, 2.509)
5	-0.060	(-0.309,0.189)	1.718	(0.674,2.762)	1.718	(0.674,2.762)	-0.078	(-0.869, 0.357)	1.751	(0.612, 2.693)
6	0.004	(-0.156,0.164)	1.650	(0.594,2.707)	1.650	(0.594,2.707)	0.026	(-0.460, 0.460)	1.568	(0.549, 2.549)
7	-0.210	(-0.815,0.395)	1.863	(0.868,2.859)	1.863	(0.868,2.859)	0.014	(-0.283, 0.306)	1.649	(0.699, 2.610)
8	-0.360	(-1.025,0.305)	2.016	(1.070,2.962)	2.016	(1.070,2.962)	0.032	(-6.682, 7.932)	1.525	(0.341, 2.687)
9	0.001	(-0.080,0.082)	1.656	(0.619,2.694)	1.656	(0.619,2.694)	-0.004	(-0.201, 0.224)	1.646	(0.715, 2.530)
10	-0.006	(-0.010,0.089)	1.659	(0.616,2.702)	1.659	(0.616,2.702)	-0.001	(-0.259, 0.248)	1.620	(0.465, 3.010)
11	-0.317	(-1.222,0.587)	1.970	(0.861, 3.080)	1.970	(0.861, 3.080)	0.205	(-0.389, 1.495)	1.438	(0.526, 2.371)
12	-0.059	(-0.305,0.188)	1.706	(0.666,2.746)	1.706	(0.666,2.746)	-0.060	(-0.706, 0.633)	1.398	(-0.616, 3.429)
13	-0.332	(-0.996,0.332)	1.973	(1.027,2.919)	1.973	(1.027,2.919)	0.003	(-5.980, 4.960)	1.660	(0.657, 2.540)
14	0.002	(-0.152,0.155)	1.650	(0.620,2.681)	1.650	(0.620,2.681)	-0.001	(-0.167, 0.174)	1.730	(0.624, 2.846)
15	0.139	(-0.236,0.513)	1.518	(0.461,2.574)	1.518	(0.461,2.574)	0.095	(-1.644, 1.734)	1.439	(0.044, 2.557)
16	0.089	(-0.224,0.402)	1.565	(0.508,2.622)	1.565	(0.508,2.622)	0.073	(-0.610, 0.915)	1.406	(0.036, 3.104)
17	-0.009	(-0.319, 0.300)	1.660	(0.606, 2.714)	1.660	(0.606, 2.714)	0.247	(-15.629, 22.704)	1.653	(0.616, 2.630)
18	-0.274	(-0.911,0.363)	1.926	(0.921, 2.930)	1.926	(0.921, 2.930)	0.030	(-0.135, 0.325)	1.656	(0.864, 2.508)
19	-0.052	(-0.302,0.197)	1.691	(0.646,2.736)	1.691	(0.646,2.736)	-0.058	(-2.533, 2.534)	1.670	(0.632, 2.853)
20	0.346	(-0.361,1.053)	1.309	(0.331,2.286)	1.309	(0.331,2.286)	0.012	(-0.242, 0.342)	1.569	(-0.273, 3.456)
21	0.060	(-0.199,0.318)	1.599	(0.545,2.651)	1.599	(0.545,2.651)	0.069	(-1.187, 1.541)	1.688	(-0.536, 3.671)
22	-0.001	(-0.003,0.028)	1.655	(0.614,2.697)	1.655	(0.614,2.697)	0.001	(-0.187, 0.157)	1.663	(0.757, 2.724)
23	0.128	(-0.285,0.540)	1.529	(0.440,2.621)	1.529	(0.440,2.621)	0.132	(-1.634, 1.658)	1.519	(-0.296, 3.460)
24	-0.289	(-0.963,0.385)	1.943	(0.877,3.008)	1.943	(0.877,3.008)	0.032	(-0.629, 0.595)	1.673	(0.733, 2.663)
25	0.280	(-0.334,0.893)	1.376	(0.451,2.302)	1.376	(0.451,2.302)	0.031	(-0.131, 0.313)	1.668	(0.542, 2.648)
26	-0.100	(-0.524,0.324)	1.761	(0.693,2.828)	1.761	(0.693,2.828)	-0.099	(-0.671, 0.210)	1.794	(0.903, 2.719)
27	0.573	(-0.246,1.393)	1.081	(0.303,1.858)	1.081	(0.303,1.858)	0.017	(-0.168, 0.231)	1.520	(0.901, 2.171)
28	-0.589	(-1.426,0.248)	2.245	(1.210,3.280)	2.245	(1.210,3.280)	0.008	(-0.307, 0.301)	1.956	(0.641, 3.248)
29	0.382	(-0.310,1.074)	1.273	(0.508,2.038)	1.273	(0.508,2.038)	-0.001	(-0.163, 0.209)	1.666	(0.607, 2.717)
30	0.001	(-0.106,0.108)	1.656	(0.618,2.693)	1.656	(0.618,2.693)	-0.002	(-0.452, 0.255)	1.620	(0.621, 2.660)
31	0.272	(-0.333,0.877)	1.379	(0.414,2.345)	1.379	(0.414,2.345)	0.127	(-0.307, 0.797)	1.508	(0.465, 2.659)
32	-0.005.	(-0.102,0.092)	1.662	(0.616,2.707)	1.662	(0.616,2.707)	-0.028	(-0.464, 0.207)	1.694	(0.770, 2.735)
33	0.128	(-0.235,0.490)	1.528	(0.471,2.585)	1.528	(0.471,2.585)	0.112	(-0.235, 0.526)	1.470	(0.322, 2.508)
34	-0.081	(-0.811,0.647)	1.740	(0.549,2.931)	1.740	(0.549,2.931)	0.070	(-0.524, 1.225)	1.600	(0.691, 2.528)
35	0.433	(-0.301,1.167)	1.239	(0.466,2.012)	1.239	(0.466,2.012)	0.021	(-5.123, 4.802)	1.614	(0.760, 2.450)

Future directions

This is a novel approach that can disentangle the mediation effects for zero-inflated mediators.

Some future directions:

- Non-parametric mediation effect.
- Evaluate all mediators in one model.
- Sensitivity analysis

Acknowledgment

Postdoc: Janaka S.S. Liyanage. Grant: R01GM123014

University of Florida: Susmita Datta, Christian Jobin, Raad Gharaibeh

Dartmouth: James O'Malley